

***Flavobacterium heparinum* EXPRESSION SYSTEM**

Field of the Invention

5

The present invention is in the area of a novel procaryotic expression system and, in particular, a *Flavobacterium heparinum* host system for the expression of gene products.

10 Background of the Invention

Genetic engineering has made it possible to produce large amounts of heterologous proteins or polypeptides in bacterial cells by means of recombinant expression systems. The expressed heterologous proteins may be mammalian, other 15 eukaryotic, viral, bacterial, cyanobacterial, archeabacterial or of synthetic origin.

The bacterial host cell of choice for expression of recombinant DNA has long been *Escherichia coli*. However, this microorganism has many disadvantages as a host cell including the inability to secrete recombinantly produced proteins, the 20 precipitation of highly expressed recombinantly produced proteins into inclusion bodies within the cell, the inability to glycosylate recombinantly produced proteins and the difficulties faced when purifying a recombinantly produced protein from *E. coli*.

25 The DNA sequences and methods for the expression of homologous and heterologous proteins in *Flavobacterium heparinum*, a Gram negative, non-pathogenic soil bacterium, have not been investigated. The ability to express heterologous and homologous polypeptides and proteins in *F. heparinum* would be desirable as the microorganism naturally produces glycosylated proteins. Thus, *F. heparinum* would 30 be useful as a host cell for expression of recombinant DNA encoding mammalian proteins since they are often glycosylated. In addition, *F. heparinum* naturally produces low levels of the glycosaminoglycan degrading enzymes that act upon heparin and other glycosaminoglycans. The heparinase enzymes have significant medical utility and the production of these enzymes from their natural source would be 35 desirable.

It is an object of the invention to provide nucleic acids and expression systems for the expression of recombinant DNA molecules in *F. heparinum*. Another object of the invention is to provide methods for use of *F. heparinum* as a host cell for the expression of recombinant DNA molecules.

5

Summary of the Invention

It has been discovered that *F. heparinum* is useful as a host cell for the expression of recombinant DNA sequences. One aspect of the present invention provides a method which achieves expression of homologous and heterologous genes in *F. heparinum* and the coupling of expression to secretion to produce a biologically active polypeptide or protein. Further, an expression system for cloned gene sequences comprising the host, nucleotide sequences encoding a polypeptide or protein of interest, and a vector. In a preferred embodiment of the system, the vector for expression of homologous and heterologous sequences comprises a functional origin of replication (*OriC*) from *F. heparinum* or another functional origin of replication, replication (*rep*) genes which direct the replication of the vector, a promoter derived from a gene endogenous to the *F. heparinum* host. In another preferred embodiment, the vector for expression of homologous and heterologous sequences comprises a chromosomal integration sequence. In a further embodiment, the expression vector further comprises selective markers and/or a regulated promoter for the expression of desired gene sequences.

The overexpression of homologous and heterologous cloned genes in the *F. heparinum* host cell, and the suitability of *F. heparinum* for the expression of recombinant proteins is demonstrated by the expression of the enzymes, heparinase I, heparinase II, heparinase III, chondroitinase AC, and chondroitinase B from *F. heparinum*; dhfrII and β -galactosidase encoding genes from *E. coli*.

30 Brief Description of the Drawings

Fig. 1 is a schematic diagram of plasmid pIBXF2 which is derived from plasmid pBBCm.

35

Fig. 2 is a schematic diagram of the integration vector pIBXF1.

Detailed Description of the Invention

The invention is the use of the bacterium *Flavobacterium heparinum* for the expression of homologous and heterologous genes. The invention provides nucleic acids and expression systems for the expression of recombinant DNA molecules in *F. heparinum*. The invention further provides methods for introducing vectors and other recombinant DNA into *F. heparinum*.

10 Definitions

“Biologically active.” The term “biologically active” as used herein in relation to homologous or heterologous polypeptides or proteins expressed in the *F. heparinum* host cell denotes that the polypeptide or protein is produced in an appropriate 15 conformation for obtaining biological activity, rather than a misfolded, aggregated and insoluble form which requires special denaturation and renaturation measures to achieve an appropriate conformation to any substantial extent. The term therefore includes polypeptide precursors of biologically active proteins, which are not themselves biologically active, but which can be readily converted into their 20 biologically active forms.

“Foreign” or “heterologous” or “non-bacterial”; “native” or “homologous”. A “foreign” or “heterologous” polypeptide is a peptide which is not normally found in a host cell of a particular species. The nucleic acid encoding such a polypeptide is also 25 referred to as “foreign” or “heterologous.” For example, insulin-like growth factor (IGF), insulin-like growth factor binding protein (IGFBP), and transforming growth factor-Beta (TGF- β) are native to mammalian cells and human rhinovirus 3C protease is native to viruses and virally-infected mammalian cells, but these proteins are foreign or heterologous to *F. heparinum*. A “non-bacterial protein” is a protein or polypeptide 30 which is not naturally found in a bacterial cell. Non-bacterial proteins include viral and eukaryotic proteins. Non-bacterial, foreign, or heterologous proteins may also be fusions between non-bacterial, foreign, or heterologous proteins and other proteins or polypeptides. For the embodiments encompassed by this invention, both 35 “heterologous protein” and “non-bacterial protein” may be expressed. As disclosed herein, genes encoding heterologous or non-bacterial proteins of interest do not

contain promoters functional in the host cell. The genes must be linked to a separate promoter that is functional in the host cell in order to be expressed. A "native" or "homologous" polypeptide or DNA sequence, by contrast, is commonly found in the host cell. A promoter or other sequence effecting, for example, the transcription or 5 translation of a gene is also considered "homologous" if it is functional in the host cell. For example, a *hepA* gene (coding for heparinase I) is considered "homologous" to a *F. heparinum* host cell, since, the encoded protein, heparinase I, is naturally produced by the cell.

10 "Genes encoding heterologous, foreign or non-bacterial proteins". "Genes encoding heterologous, foreign or non-bacterial proteins" contain all of the genetic elements necessary for the expression of the heterologous, foreign or non-bacterial protein with the exception of a promoter functional in the host cell. These genes encompass recombinant genes which may include genetic elements native to the host 15 cell. Further, the coding regions of these genes may optionally be optimized for the codon usage of the host cell.

20 "Operably linked". A nucleic acid sequence is operably linked when it is placed into a functional relationship with another nucleic acid sequence. For example, a promoter is operably linked to a coding sequence if the promoter affects its transcription or expression. Generally, DNA sequences which are operably linked are contiguous and, where necessary, in reading frame.

25 "Recombinant". A "recombinant" nucleic acid is one which is made by the joining of two otherwise separated segments of nucleic acid sequence in vitro or by chemical synthesis.

30 "Vector". A "vector" is a carrier DNA molecule into which a nucleic acid sequence can be inserted for introduction into a host cell, and the inserted nucleic acid sequences may be expressed. A vector may replicate autonomously in the host cell or be integrated into the host cell genome.

35 As detailed below, there are a number of methods and variables to consider in the construction of an expression system using *F. heparinum* as the host. Depending on the protein to be expressed and the degree of regulation and quality of expression

desired, the system can consist of as little as the host and an expression vector containing transcription and translation signals and the gene to be expressed. To control the rate and extent of expression, the gene of interest may be placed under the control of an inducible promoter or a constitutively expressed promoter, or the 5 promoter can be placed under the control of a repressor or stimulatory protein.

The first consideration is to determine the polypeptide or protein to be expressed, and to isolate a sequence encoding some or all of the protein. Methods for the isolation of protein-encoding sequences are well known to those of ordinary skill 10 in the art of genetic engineering. Examples of the isolation and characterization of the exemplary proteins are provided below.

The second consideration is the selection and/or construction of a vector, such as a plasmid which provides for homologous recombination or other recombinant 15 DNA having an origin of replication, and transposable elements, including bacteriophage systems, transposons and insertion sequences. The plasmid system can be a broad-host or narrow-host plasmid. General techniques for nucleic acid manipulation useful for the practice of the claimed invention are described generally, for example, in Sambrook et al., "Molecular Cloning: A Laboratory Manual", Vols. 1- 20 3 (Cold Spring Harbor Laboratory Press, 2 ed., 1989) or in Ausubel et al., Current Protocols in Molecular Biology (Green Publishing and Wiley-Interscience: New York, 1987).

Useful vectors will preferentially be a modified broad-host plasmid which can 25 replicate in *F. heparinum*, or a vector containing a DNA fragment which can facilitate the homologous recombination. In addition, a vector will include restriction enzyme sites for insertion of foreign DNA and one or more selectable markers. Commonly used selectable markers are genes that confer resistance to antibiotics such 30 as ampicillin, tetracycline, chloramphenicol, erythromycin, and trimethoprim. Additional selectable markers include genes which encode for heavy metal resistance, such as mercuric reductase, or a nutritional requirement factor, such as an amino acid requirement factor. A preferred selectable marker is the trimethoprim-resistance dihydrofolate reductase II gene. The selective marker preferably is regulated by a regulatory region from *F. heparinum*. Useful regulatory regions include the 35 heparinase I gene regulatory region. Other useful plasmids, and the genetic

engineering methods for inserting genes of interest into the plasmid will be apparent to those of ordinary skill in the art.

5 Optionally, the desired protein encoded by the recombinant DNA can be genetically engineered to be secreted by the addition of secretion signals and processing signals in operative association with the nucleic acid sequence encoding the desired protein. The molecular techniques to engineer a vector to include secretion and processing signals are well known to those of ordinary skill in the art.

10 Once the nucleotide sequence and the vector have been selected, sequences for the expression, regulation of expression, and post-translational characteristics of the expression protein can be isolated and inserted into the vector. A number of promoters are useful in the present invention, including the promoters for the homologous genes *hepA* and *lysA* and heterologous promoters.

15 A variety of methods for introducing nucleic acids into host cells are known in the art, including, but not limited to, electroporation; transfection employing calcium chloride, rubidium chloride calcium phosphate; DEAE-dextran, or other substances; microprojectile bombardment; lipofection; and infection (where the vector is an 20 infectious agent, such as a retroviral genome). See generally, Sambrook et al., *supra* and Ausubel et al., *supra*. Methods for bacterial transformation by plasmid DNA have been developed to optimize the efficiency of DNA uptake. Generally, the bacterial cells are treated to render them competent to incorporate plasmid DNA, and then once 25 inside, the selectable marker allows the transformed cells to survive in the presence of the selected medium.

Some plasmids can bring about their own transfer by the conjugation process, as is well understood in the art. Briefly, a cell extrudes a sex pilus, which is coded for by plasmid genes, which adheres to a neighboring cell, and the cells are then drawn 30 into direct contact. The plasmid then undergoes transfer replication and one parental strand passes to the recipient cell and one strand remains in the donor cell. Complementary strands are synthesized both in the donor and recipient cells simultaneously with the transfer. The daughter molecules are circularized by ligase 35 action immediately after the transfer replication is complete. The mating cells then break apart.

Other methods can achieve either or both of these results by transfer and integration of DNA into a suitable site on the chromosome. The transferred DNA comprises a DNA fragment encoding a selectable marker and a homologous DNA sequence from the host. Optionally, a transferred DNA may also comprise, in an operable linkage to the sequence encoding the homologous DNA, transcription and translation initiation regulatory sequences and expression control sequences, which may include a promoter, an enhancer and necessary processing information sites, such as ribosome-binding sites, and mRNA stabilizing sequences, as well as any necessary secretion signals, where appropriate, which allow the expressed polypeptide or protein to cross and/or lodge in cell membranes, and thus attain its functional topology, or be secreted from the cell. Alternatively, integration of a gene of interest near a promoter on the *F. heparinum* chromosome can be designed to result in a operable linkage (for example, by integrating a transferred DNA into an operon on the chromosome). The site of integration or sequences adjacent to the site of integration may facilitate amplification (e.g. where the site is located in a transposable element, by providing duplicate DNA sequences, or even by providing a region of DNA sequence homologous to a portion of the chromosomal transfer DNA, thus providing duplicate DNA sequences).

20

Examples

The following examples are provided to illustrate the invention, but are not intended to, and should not be construed to, in any way limit the scope of the invention as defined by the claims which follow.

25

The Examples do not include detailed descriptions for conventional methods employed in the construction of plasmids, the insertion of genes encoding polypeptides into such vectors or the introduction of the resulting plasmids into the *F. heparinum* hosts. The Examples also do not include detailed description for conventional methods employed to assay and identify the polypeptides produced by such host vector systems. Such methods are well known to those of ordinary skill in the art and are described in numerous publications.

Example 1. *F. heparinum* Plasmid Systems:

5 pBBR1Cm is a derivative of the plasmid pBBR1 with a gene coding for chloramphenicol resistance cloned in. It is a cryptic plasmid isolated from *Bordetella bronchiseptica* S87. It is further a broad-host-range plasmid known to replicate in *E. coli*.

pIBXF2 has the following characteristics:

- (1) a *bla* gene coding for β lactamase (conferring ampicillin resistance) – functional in *E. coli*;
- 10 (2) an ORI T region – allows for the mobilization of the plasmid into a new host (a *Sau3* AI 307bp fragment isolated from plasmid RP4);
- (3) a *dhfrII* gene coding for dihydrofolate reductase type II (conferring trimethoprim resistance) – cloned from plasmid R751 and placed under the control of 15 the *hepA* promoter – functional in *F. heparinum*;
- (4) a *hepA* promoter – an 830 bp fragment 5' of the *hepA* gene – induced by heparin;
- 20 (5) a unique restriction site (for example, a *Xba1* site) – site for insertion of all homologous and heterologous genes;
- (6) a *rep* gene – involved in the replication of the plasmid; and
- (7) a putative ORI – functional in both *E. coli* and *F. heparinum* – A-T rich region common to ORI C.

25 pIBXF2 was constructed, using methods known in the art, fusing portions of pIBXF1 and pBBR1Cm. DNA amplification was used to generate a fragment from pIBXF1 which included the region from the ampicillin resistance marker to the *Xba1* site (see Figure 2) – the ColEI ORI was not included. This fragment was inserted into pBBR1Cm within the gene encoding chloramphenicol resistance (not shown) disrupting it.

30 Once pIBXF2 was constructed, homologous genes *cslA* ad *cslB* were separately cloned into the plasmid. The homologous genes were under the control of the *hepA* gene promoter.

35 The following heterologous gene has been cloned into the vector: *dhfrII*.

Example 2. *F. heparinum* Integration Vector

5 The integration vector, pIBXF1, was created by the addition of various elements to the parent plasmid, pUC13T .

The following elements are present on the parent plasmid:

10 (1) *bla* gene coding for β -lactamase (conferring ampicillin resistance) – functional in *E. coli*.
(2) ColEI ORI – origin of replication – functional in *E. coli*.
(3) ORI T region – allows for the mobilization of the plasmid into a new host (a *Sau3* AI 307 bp fragment isolated from plasmid RP4).

15 The following elements were introduced into the pUC13T plasmid to confer integrative and selective functions in *F. heparinum*.

20 (1) 10kb *HindIII* fragment – a 10 kb fragment randomly selected from genomic *F. heparinum* DNA which was digested with *HindIII*, and cloned into pUC13T.
(2) *dhfrII* gene coding for dihydrofolate reductase type II (conferring trimethoprim resistance) – cloned from plasmid R751 and placed under the control of the *hepA* promoter – functional only in *F. heparinum*.
(3) *hepA* promoter – an 830 bp fragment 5' of the *hepA* gene – induced by heparin.

25 Once the *F. heparinum* integration vector was constructed (see Figure 1), homologous genes *cslA*, *cslB*, *hepA*, *hepB*, and *hepC* were separately cloned into the vector. The homologous genes were placed under the control of the *hepA* promoter.

30 The heterologous genes *dhfrII* and *lacZ* were also cloned into the vector.

Example 3. Introduction of Recombinant DNA into *F. heparinum*

Conjugation:

5 Plasmids were introduced into the *E. coli* strain S17-1. S17-1 is a mobilizing strain which provides, in *trans*, the genes needed for conjugative DNA transfer.

10 Overnight cultures of both *F. heparinum* and *E. coli* S17-1 containing the plasmid DNA were subcultured to an OD₆₀₀ of 0.1 and grown to a mid-log phase of OD₆₀₀ of 0.5. *F. heparinum* was grown in MG media supplemented with 1% heparin, 0.02% methionine/histidine, and 2 mM MgSO₄ at 23°C. *E. coli* was grown in Lauria-broth supplemented with 50 µg/mL ampicillin and 10 µg/mL trimethoprim at 37°C.

15 1.75 mL of *F. heparinum* and 100 µL of *E. coli* were combined in a 2 mL eppendorf tube and spun at 2500 g for 3 minutes. The cell pellet was resuspended in 100 µL MG and spread over a 0.45 mm membrane (Millipore, catalog number HAWG 025 00) which was placed on an agar petri dish (preheated for 30 minutes at 30°C) containing LB/MG medium supplemented with 1% heparin, and 50 µg/mL ampicillin and incubated overnight at 30°C. As a negative control, 1.75 mL of *F. heparinum* was 20 treated as above.

25 Bacteria which grew on the membrane were harvested by placing the membrane into 1 mL of MG and vortexing until bacteria could no longer be seen on the membrane. The bacteria were then placed on selective agar plates (called MGHKT) consisting of MG medium supplemented with 1% heparin, 0.02% methionine/histidine, 2 mM MgSO₄, 100 µg/mL gentamicin, 100 µg/mL kanamycin, and 0.6 mg/mL trimethoprim. The cells were plated at the following ratios: 1/5, 3/10, and ½ of the total harvested cell volume. The plates were then incubated at 23°C for approximately 5 to 7 days until colonies were visible .

30

Electroporation:

35 *F. heparinum* cells were made electro-competent according to the method described by Sambrook et al., *supra*. The DNA and 40 µL of electro-competent *F. heparinum* (thawed on ice) were combined in a pre-chilled (on ice) 1.5 mL eppendorf

5 tube and allowed to sit on ice for approximately 1 minute. This mixture was then placed in a pre-chilled (on ice) electroporation cuvette with a 0.2 cm electrode gap (Bio-Rad Laboratories, catalogue number 165-2086). The cuvette was placed in a Gene-Pulser electroporation apparatus (Bio-Rad Laboratories, catalogue number 165-2098). A resistance of 400 Ω with a capacitance of 25 μ FD and a voltage of 2.5 kV was applied to the cells for a field strength of 12.5 kV/cm. The cells were then 10 immediately resuspended in 1 mL Lauria-broth supplemented with 1 % heparin and incubated at 30°C for 4-6 hours. The cells were then microfuged at 2500 g for 3 minutes and plated on MGHKT selective plates. The plates were incubated at 23°C for 5-7 days until colonies became visible .

Example 4. Expression of Proteins in *F. heparinum*

15 The following proteins have been expressed in *F. heparinum* using the integration vector, pIBXF1, containing the genes described below:

20

- *cslA* gene with GenBank accession number U27583; coding for the chondroitinase AC protein.
- *cslB* gene with GenBank accession number U27584; coding for the chondroitinase B protein.
- *hepB* gene with GenBank accession number U27585; coding for the heparinase II protein.
- *hepC* gene with GenBank accession number U27586; coding for the heparinase III protein.
- *hepA* gene with GenBank accession number L12534; coding for the heparinase I protein.
- *lacZ* gene with GenBank accession number L08936 (plasmid pMC 1871).

25 The proteins which have been expressed in *F. heparinum* using the plasmid system, pIBXF2, are those of the following genes' products:

30

- *cslA* gene with GenBank accession number U27583; coding for the chondroitinase AC protein.
- *cslB* gene with GenBank accession number U27584; coding for the chondroitinase B protein.

35

Also, the trimethoprim resistance gene, *dhfrII*, from plasmid R751 was expressed in both systems.

All genes were placed under control of the *hepA* promoter, with the exception of *lacZ*, which was placed under the control of the *F. heparinum lysA* promoter. In the case of the genes coding for the chondroitinases and heparinases, their respective leader peptides were included.

The GAG lyase proteins mentioned above, with the exception of heparinase III, were glycosylated, as were the native proteins isolated from *F. heparinum*. These proteins were further characterized by SDS-PAGE analysis for molecular weight, Western analysis using polyclonal antibodies specific for non-glycosylated moieties of the various proteins, and by enzymatic assays in which the appropriate substrate was degraded by the enzyme. In all cases, the transconjugate protein products behaved in the same manner as their native counterparts with the exception that greater specific activities (IU ml⁻¹OD₆₀₀⁻¹) were seen for the transconjugate proteins.

The *dhfrII* gene protein product was characterized by the ability of the transconjugate colony to be resistant to high levels of trimethoprim, i.e. 0.6 mg/mL.

20

Example 5. Culturing of *F. heparinum*.

F. heparinum, ATCC 13125, produces heparinases I, II, and III, and chondroitinases AC and B. It is also referred to as *Cytophaga heparina* (Ref. Christensen, 1980, Int. J. Syst. Bacteriol. 30:473-475); *Sphingobacterium heparinum* (Ref. Takeuchi and Yokota, 1992, J. Gen. Appl. Microbiol. 38:465-482); or *Pedobacter heparinus* (Ref. Steyn et al., 1998, Int. J. Syst. Bacteriol., 48:165-177).

F. heparinum cells can be cultured in an ordinary medium containing a carbon source, nitrogen source, and inorganic salts. As a carbon source, any of those which can be assimilated can be used, for example, D-glucose, maltose, D-xylose, sucrose, heparin, citrate, succinate, glutamate, tryptone, or peptone are typical examples. As a nitrogen source, various conventional materials can be used, for example, yeast extract, peptone, meat extract, amino acids or organic nitrogen such as ammonia sulfate or ammonium chloride. In addition, various salts may be added, for example,

inorganic salts such as magnesium sulfate, magnesium chloride, potassium chloride, calcium chloride, potassium phosphate, and sodium phosphate. Vitamins and other growth factors are not essential but may be added, for example, folic acid, calcium pantothenate, biotin, riboflavin, and thiamine. It is further possible to add a gelling 5 agent when desired, such as agar, gelatin, and gellan gum. The medium is adjusted to a final pH of 5.5 to 8.5, preferably pH 7.0, by addition of an acid or base as appropriate.

As a specific example of a suitable defined medium, a liquid medium containing (per liter) 6 g K_2HPO_4 , 3 g $KH_2PO_4 \cdot H_2O$, 1 g $NaCl$, 2 g NH_4Cl , 2 mM 10 $MgSO_4 \cdot 7H_2O$, 0.4 g L-histidine, 0.4 g L-methionine, 16 g glucose, 2 g heparin, 10^{-4} mM each of $NaMoO_4 \cdot 2H_2O$, $CoCl_2 \cdot 6H_2O$, $MnSO_4 \cdot 7H_2O$, $CuSO_4 \cdot 5H_2O$, $FeSO_4 \cdot 7H_2O$ and $CaCl_2$. The organism can be stored for up to two weeks on agar plates containing 1% Disco agar in defined medium, containing 4 g heparin/L as the sole carbon source, or the organism can be stored indefinitely in 10% glycerol at $-80^{\circ}C$.

15

Example 6. Purification of Proteins from *F. heparinum*.

Methods which have been used to extract proteins from the periplasmic space of Gram negative bacteria include osmotic shock treatment as the initial step. 20 Typically these procedures include an initial disruption in osmotically stabilizing medium followed by selective release in non-stabilizing medium. The composition of these media (pH, protective agent) and the disruption methods used (chloroform, lysozyme, EDTA, sonication) vary among specific procedures reported.

25 *F. heparinum* cells can be concentrated by ultrafiltration, and subjected to osmotic shock treatment to release proteins from the periplasmic space. Disruption of the cell envelope can be induced by exposing the cells to an osmotically stabilized medium (containing 20% sucrose) with or without EDTA, followed by an initial release of periplasmic material into a non-stabilized medium (10 mM phosphate, at a 30 pH between 6.0 and 8.6) with a subsequent release into a second non-stabilized medium (10 mM phosphate, 150 mM sodium chloride, at a pH between 6.0 and 8.6). This three step procedure can provide an initial fraction of the proteins in the periplasmic space. As will be apparent to one skilled in the art, the purification method used will depend on the identity of the foreign protein.